

## Universality of filamentous aggregation phenomena

Thomas C. T. Michaels,<sup>1,2</sup> Alexander J. Dear,<sup>1</sup> and Tuomas P. J. Knowles<sup>1,3,\*</sup>

<sup>1</sup>*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*

<sup>2</sup>*Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*

<sup>3</sup>*Cavendish Laboratory, Department of Physics, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, United Kingdom*



(Received 29 April 2019; published 27 June 2019)

We use perturbative renormalization group theory to study the kinetics of protein aggregation phenomena in a unified manner across multiple timescales. Using this approach, we find that, irrespective of the specific molecular details or experimental conditions, filamentous assembly systems display universal behavior in time. Moreover, we show that the universality classes for protein aggregation correspond to simple autocatalytic processes and that the diversity of behavior in these systems is determined solely by the reaction order for secondary nucleation with respect to the protein concentration, which labels all possible universality classes. We validate these predictions on experimental data for the aggregation of several different proteins at several different initial concentrations, which by appropriate coordinate transformations we are able to collapse onto universal kinetic growth curves. These results establish the power of the perturbative renormalization group in distilling the ultimately simple temporal behavior of complex protein aggregation systems, creating the possibility to study the kinetics of general self-assembly phenomena in a unified fashion.

DOI: [10.1103/PhysRevE.99.062415](https://doi.org/10.1103/PhysRevE.99.062415)

### I. INTRODUCTION

The formation of protein filaments is a ubiquitous example of a self-assembly phenomenon that has fundamental implications for biology [1,2], medicine [3–5], and materials science [6–8]. For instance, biofilaments of actin and tubulin are key components of the cellular cytoskeleton, which is implicated in cell shape regulation and cell division [1,2], while a particular class of protein filaments, known as amyloids, are associated with over 50 medical disorders, including Alzheimer's and Parkinson's diseases [3–5]. One of the most intriguing yet least understood aspects of protein filament formation is its generality, i.e., the fact that many different proteins, with unrelated sequence or fold, are able to self-assemble into such filamentous structures [9,10]. A central question in this area is thus to establish whether certain aspects of the aggregation process are universal, i.e., are independent of the molecular specifics and are thus conserved across different protein systems.

Historically, renormalization group (RG) theory approaches [11] have emerged as the fundamental tool for explaining how microscopically different systems can display universal features at the macroscopic level in areas ranging from condensed matter physics to particle physics. Asymptotic analysis seeks to obtain the simplified intermediate asymptotic behavior of a system of equations [12], for instance, through the identification and solution of an amplitude equation [13] or the discovery of a slow manifold [14]. Since asymptotic analysis involves identifying a reduced lower-dimensional description of the behavior of a physical system, we might expect disparate systems to display universal

features asymptotically, just as in classical RG problems. In this context, Goldenfeld and co-workers [14–16] demonstrated a deep connection between RG theory and asymptotic analysis of singular perturbation problems and developed a highly general perturbative RG approach to asymptotic analysis that was shown to be superior in accuracy to many standard methods. In this paper we apply perturbative RG theory to protein aggregation kinetics and find that the macroscopic behavior of these phenomena is universal in time. In particular, we find that disparate protein aggregation systems can be categorized into universality classes that are indexed by a single scaling exponent which is set solely by the dependence of fibril self-replication on the supply of monomeric protein. Within these universality classes, kinetic profiles for the aggregation of proteins with a wide range of distinct structures and biological roles collapse onto a single universal curve, irrespective of specific experimental conditions. The experimentally reported scaling behavior of protein aggregation [17] emerges as a direct consequence of this universality.

### II. PERTURBATIVE RG APPROACH TO PROTEIN AGGREGATION

#### A. Kinetic equations for protein filament formation

Protein aggregation is the result of an interplay between multiple molecular events (Fig. 1). Generally, protein aggregation is initiated by a primary nucleation step, whereby monomers in solution come together spontaneously to form the smallest stable aggregate, followed by filament growth through elongation at one or both ends of the aggregates [1]. For many protein systems, including the aggregation of disease-related species such as Alzheimer's amyloid- $\beta$  peptide, aggregation is accelerated by secondary processes that, unlike primary nucleation, depend on the actual fibril

\*tpjk2@cam.ac.uk

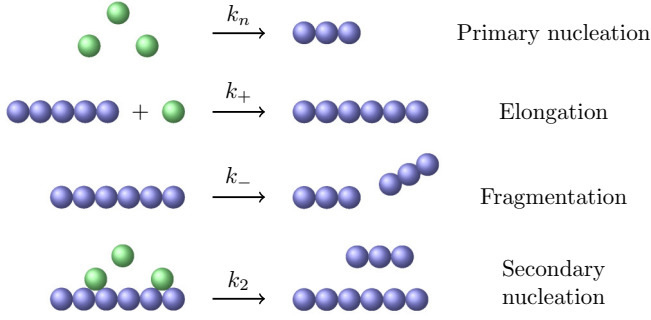


FIG. 1. Schematic representation of the well-established elementary steps of protein filament formation and definition of the associated rate constants.

concentration and hence lead to an autocatalytic multiplication of the fibrillar structures; key examples of such secondary processes are filament fragmentation [17–22], lateral branching [23–25], and surface-catalyzed secondary nucleation [26–31]. The interplay between these different microscopic steps of aggregation can be captured by means of a master equation approach [32,33], which tracks the time evolution of the population of aggregates of different sizes. In the context of protein aggregation, however, the experimentally accessible information about the aggregation process is commonly encoded in the time evolution of three coarse-grained fields: the number and mass concentrations of fibrils, denoted by  $P(t)$  and  $M(t)$ , and the free monomer concentration  $m(t)$ . The time evolution of these experimentally observable quantities, which correspond to the principal moments of the filament distribution, can be derived explicitly in terms of the underlying mechanisms of aggregation from the non-linear master equation describing the time evolution of the entire distribution of aggregate sizes [33]; for an aggregating system evolving through primary and secondary nucleation pathways, this procedure results in the general set of kinetic equations [33]

$$\frac{dm(t)}{dt} = -2k_+m(t)P(t) = -\frac{dM(t)}{dt}, \quad (1a)$$

$$\frac{dP(t)}{dt} = k_n m(t)^{n_c} + k_2 m(t)^{n_2} [m_{\text{tot}} - m(t)], \quad (1b)$$

where  $m_{\text{tot}} = M(t) + m(t)$  is the conserved total monomer concentration. Equation (1a) describes the consumption of monomers and consequent buildup of aggregate mass through the elongation of existing filaments with rate constant  $k_+$ .<sup>1</sup>

<sup>1</sup>We have implicitly assumed that both ends of a filament elongate with the same rate constant  $k_+$ , thus explaining the factor of 2 in Eq. (1a). Our framework can however be generalized in a straightforward manner to account for a different elongation rate at each end. Moreover, in Eq. (1a) we have neglected sink terms for the change in monomer mass concentration due to primary and secondary nucleation events. These terms can be neglected relative to growth at the filament ends, since for most known protein aggregation processes the ratios of the nucleation rates to the rate of growth are  $v_1 = k_n m_{\text{tot}}^{n_c-2}/2k_+ \ll 1$  and  $v_2 = k_2 m_{\text{tot}}^{n_2-1}/2k_+ \ll 1$ . Indeed, the

Equation (1b) describes the rate of formation of new fibrils through primary and secondary nucleation, with the rate constants for these processes being  $k_n$  and  $k_2$ . The key difference between primary and secondary nucleation processes is that the rate of the latter depends on the existing fibril population. The dependence of the rates of primary and secondary nucleation on the concentration of precursor monomer is captured by means of the reaction orders  $n_c$  and  $n_2$ , respectively. Various possible secondary nucleation processes correspond to different values of the reaction order  $n_2$ : Filament fragmentation, which is independent of monomer concentration, corresponds to  $n_2 = 0$ , lateral branching corresponds to a linear dependence on the monomer concentration ( $n_2 = 1$ ), and surface-catalyzed secondary nucleation is usually described by higher reaction orders ( $n_2 \geq 2$ ). It is important to note that both primary and secondary nucleation of new filaments are believed to be nonclassical, multistep nucleation processes featuring metastable oligomeric intermediate species. These species are likely micellar in nature, possess low  $\beta$ -sheet content, and are thus structurally distinct from mature fibrils, as observed both in experiments [34] and in computer simulations [30,35–37]. The rate law employed in our model equation (1b) is a coarse-grained description of these nonclassical nucleation processes, which can be justified from an explicit consideration of the free-energy landscape in aggregate size and structural space (see [37,38] for further details). This calculation shows that the exponents  $n_c$  and  $n_2$ , which represent the reaction orders of primary and secondary nucleation with respect to the free-monomer concentration, are not related to the overall physical size of the nucleating aggregates as in classical nucleation theory. Instead,  $n_c$  and  $n_2$  are related to the portion of the nucleating aggregate that actively participates in the conformational conversion of oligomers to fibrils [37,38].

## B. Perturbation expansion

Equations (1) do not admit exact solutions in general [26,39] and we instead seek asymptotic solutions. The perturbative RG approach is intimately connected with asymptotic analysis [14–16]; as a first step, it is thus useful to recast Eqs. (1) into a singular perturbation problem. To do so, we introduce dimensionless monomer and aggregate number concentrations as

$$\mu(t) = \frac{m(t)}{m_{\text{tot}}}, \quad \Pi(t) = \frac{\kappa P(t)}{2k_+} \quad (2)$$

steady-state average length of filaments is  $\simeq 1/\sqrt{v_2}$  [see [32,39] or consider the ratio  $M(t)/P(t)$  from Eqs. (15) and (16), which gives the average length of aggregates]. Since filaments consist typically of several thousands of monomers, it follows that  $1/\sqrt{v_2} \gg 1$ . For systems where secondary nucleation is the dominant mechanism of formation of new aggregates, this implies that  $v_1 \ll v_2 \ll 1$ .

and a dimensionless time variable as  $t' = \kappa t$ , where  $\kappa = \sqrt{2k_+k_2m_{\text{tot}}^{n_2+1}}$ . In this manner, Eqs. (1) become

$$\frac{d\mu(t')}{dt'} = -\mu(t')\Pi(t'), \quad (3a)$$

$$\frac{d\Pi(t')}{dt'} = 2\varepsilon\mu(t')^{n_c} + \mu(t')^{n_2}[1 - \mu(t')], \quad (3b)$$

where the following dimensionless parameter emerges naturally:

$$\varepsilon = \frac{\lambda^2}{2\kappa^2}. \quad (4)$$

Here  $\lambda = \sqrt{2k_+k_2m_{\text{tot}}^{n_c}}$  and  $\kappa = \sqrt{2k_+k_2m_{\text{tot}}^{n_2+1}}$  are the characteristic rates that describe aggregate proliferation via primary [1] and secondary nucleation [17,26,39,40], respectively. Equations (3a) and (3b) can then be combined and reformulated as

$$-\frac{d^2 \ln \mu(t')}{dt'^2} = 2\varepsilon\mu(t')^{n_c} + \mu(t')^{n_2}[1 - \mu(t')]. \quad (5)$$

In many cases, secondary processes dominate over primary nucleation for the formation of new filaments, i.e.,  $\varepsilon \ll 1$ . Typical values for  $\varepsilon$  are in fact  $\varepsilon \sim 10^{-2}$  for the yeast prion Ure2p [22],  $\varepsilon \sim 10^{-3}$  for sickle-cell hemoglobin (Hsb) [26],  $\varepsilon \sim 10^{-5}$  for the amyloid- $\beta$  (A $\beta$ ) peptide [28], and  $\varepsilon \sim 10^{-7}$  for the islet amyloid polypeptide (IAPP) [27]. We have thus recast Eqs. (1) into a singular perturbation problem [Eq. (5)], where the relevant perturbation parameter emerges naturally as the ratio between the rates of primary and secondary nucleation. Consequently, a second-order perturbation series solution  $\mu(t') = \mu^{(0)}(t') + \varepsilon\mu^{(1)}(t') + \varepsilon^2\mu^{(2)}(t') + O(\varepsilon^3)$  can be found for Eq. (5) by applying the initial conditions  $\mu(0) = 1$  and  $d\mu(0)/dt' = 0$  to yield

$$\mu(t') = 1 - \varepsilon e^{t'} + \frac{\varepsilon^2}{2c} e^{2t'} + \mathcal{R}, \quad (6)$$

where  $c = 3/(2n_2 + 4)$  and  $\mathcal{R}$  denotes other terms that either are of order  $\varepsilon^3$  or vanish in comparison to the dominant terms at the respective order in  $\varepsilon$  for large  $t'$ . The perturbative solution (6) approximates the exact solution of (1) for early times, i.e., the solution to the linearized form of Eqs. (1) obtained when  $m(t) = m_{\text{tot}}$ , but it fails to capture the behavior at later times [Fig. 2(a)]. This is due to exponentially divergent terms (UV divergence), a phenomenon originally pointed out by Ferrone *et al.* [26] that has challenged the derivation of accurate expressions for the aggregation kinetics. Physically, this divergence emerges because the perturbative expansion assumes constant monomer concentration, hence leading to unbounded multiplication of aggregates through secondary nucleation and growth.

### C. Perturbative RG

To tame this divergence we impose timescale invariance on our perturbative solution. This amounts to resumming the perturbation series for  $\mu$  at the observation time  $t'$  such that its dependence on the time that has elapsed since  $\mu$  was at its initial concentration (here  $t' = 0$ ) is removed. We start by defining a new variable  $\tau = e^{\kappa t}$  to simplify the mathematics.

Following the conventional workflow of the perturbative RG [14–16], we then introduce an arbitrary past-time cutoff  $\sigma$  which we will vary between the initial time and the observation point  $\tau$  and write  $\tau = (\tau - \sigma) + \sigma$  in Eq. (6) [Fig. 2(b)]. This past-time cutoff is equivalent to the UV cutoff in conventional momentum-space RG; timescale coarse graining by rewriting  $\mu(\tau)$  in terms of  $\mu(\tau - \sigma)$  and increasing  $\sigma$  is directly equivalent to coarse graining by reducing the UV cutoff and integrating out the high-frequency degrees of freedom in momentum space. Doing so, we obtain, from (6),

$$\begin{aligned} \mu(\tau - \sigma, \sigma) &= 1 - \varepsilon(\tau - \sigma) + \frac{\varepsilon^2}{2c}(\tau - \sigma)^2\varepsilon\sigma \\ &\quad + \frac{\varepsilon^2}{2c}[\sigma^2 + 2(\tau - \sigma)\sigma] + \mathcal{R}. \end{aligned} \quad (7)$$

The next step is to renormalize  $\mu$  by multiplying Eq. (7) by a renormalization constant  $\rho(\sigma)$ , whose second-order perturbation expansion in  $\varepsilon$  is

$$\rho(\sigma) = \rho_0(\sigma) + \varepsilon\delta\rho_1(\sigma) + \varepsilon^2\delta\rho_2(\sigma). \quad (8)$$

Here  $\delta\rho_1(\sigma)$  and  $\delta\rho_2(\sigma)$  are counterterms, chosen to absorb the UV divergent terms in  $\sigma$  at the respective orders in  $\varepsilon$  [14–16]. The constant initial monomer concentration is thus effectively replaced by a running coupling which evolves with the RG scale. This yields (see the Appendix for details)

$$\delta\rho_1(\sigma) = \sigma\rho_0(\sigma), \quad \delta\rho_2(\sigma) = \sigma^2\rho_0(\sigma)\left(1 - \frac{1}{2c}\right) \quad (9)$$

and we thus arrive at the renormalized second-order expansion

$$\begin{aligned} \mu'(\tau - \sigma, \sigma) &= \rho_0(\sigma)\left\{1 - \varepsilon(\tau - \sigma) \right. \\ &\quad \left. + \varepsilon^2\left[\frac{\tau^2 - \sigma^2}{2c} - \sigma(\tau - \sigma)\right] + \mathcal{R}\right\}, \end{aligned} \quad (10)$$

where a prime indicates the renormalized solution. The renormalized solution, however, cannot depend on  $\sigma$  once the observation scale  $\tau$  is reached. We thus require [14–16]

$$\left.\frac{\partial\mu'(\tau, \sigma)}{\partial\sigma}\right|_{\sigma=\tau} = 0. \quad (11)$$

This condition gives the perturbative RG equation, which, to second order in  $\varepsilon$ , reads

$$\frac{\partial\rho_0(\tau)}{\partial\tau} = -\varepsilon\left(1 - \frac{\varepsilon\tau}{\theta}\right)\rho_0(\tau) + O(\varepsilon^3), \quad (12)$$

where we have introduced the parameter

$$\theta = \frac{c}{1 - c} = \frac{3}{2n_2 + 1}. \quad (13)$$

This procedure is entirely analogous to performing a perturbative momentum-space RG calculation in statistical physics or quantum field theory; in this analogy, the electron charge or mass is replaced by the initial monomer concentration  $\rho$  and the RG procedure yields renormalized values for this quantity at different timescales [16]. The RG equation describes a fixed point; here it is a similarity solution that no longer diverges in time. The critical points of the conventional RG are equivalent to bifurcation points in parameter space in

this dynamical RG approach and can be interpreted similarly: Minute fluctuations in the system parameters at these points give rise to qualitative changes in the system properties. The flow of the renormalization constant, or initial condition, with respect to the initial time is precisely the long-wavelength motion we seek.

In (12) we recognize the expansion of the function  $1/(1 + \varepsilon\tau/\theta) = 1 - \varepsilon\tau/\theta + O(\varepsilon^2)$  such that the solution to the second-order perturbative RG equation (12) can be found to be

$$\rho_0(\tau) = \left[1 + \frac{\varepsilon\tau}{\theta}\right]^{-\theta}. \quad (14)$$

Finally, substituting  $\rho_0(\sigma)$  into Eq. (10), increasing  $\sigma$  to  $\tau$ , and rewriting the result in terms of the original dimensional parameters, we find the following renormalized solution for the total aggregate mass:

$$\frac{M(t)}{m(0)} = 1 - \left[1 + \frac{\lambda^2}{2\kappa^2\theta} e^{\kappa t}\right]^{-\theta}. \quad (15)$$

A closed-form solution for the aggregate number concentration  $P(t)$  can be obtained from Eq. (1a) by simple differentiation,  $P(t) = -1/[2\kappa_+ m(t)] dm(t)/dt$ , as

$$\frac{P(t)}{P(\infty)} = \left[1 + \frac{2\kappa^2\theta}{\lambda^2} e^{-\kappa t}\right]^{-1}, \quad (16)$$

where  $P(\infty) = \kappa\theta/2\kappa_+$ . Figure 2(a) shows a comparison between the RG solution (15) and the numerical solution of Eqs. (1). By “removing” the UV divergence, the second-order perturbative RG solution succeeds in providing a highly accurate approximate description of aggregation kinetics when, as is usually the case,  $\varepsilon \ll 1$  [Fig. 2(a)]. We note that the obtained generalized logistic solution (15) satisfies  $M(0)/m(0) = 1 - (1 + \varepsilon/\theta)^{-\theta} \simeq \varepsilon$  (for  $\varepsilon \ll 1$ ) and yields  $M(t) = 0$  only for  $t \rightarrow -\infty$ ; the quantity  $\varepsilon m(0)$  corresponds to the critical concentration at which the term in Eq. (1b) associated with secondary nucleation becomes larger than the term relating to primary nucleation. This method fails only when secondary growth processes are not present, such as

in the growth of actin filaments without branching; in this case controlled by primary nucleation, however, the kinetic equations admit an exact solution [1].

### III. UNIVERSALITY OF FILAMENTOUS PROTEIN SELF-ASSEMBLING SYSTEMS

#### A. Self-similarity of protein aggregation

We now discuss some key consequences of our RG approach to protein aggregation. The first key prediction is that the time course of fibril mass formation has a self-similar form

$$\frac{M(t)}{m(0)} = \Phi_\theta\left(\frac{\lambda^2 e^{\kappa t}}{2\kappa^2}\right), \quad (17)$$

where

$$\Phi_\theta(x) = 1 - \left(1 + \frac{x}{\theta}\right)^{-\theta}, \quad \theta = \frac{3}{2n_2 + 1}. \quad (18)$$

This means that all instances of filamentous growth kinetics dominated by secondary processes are controlled by a single effective variable  $x = \lambda^2 e^{\kappa t}/2\kappa^2$  containing all dimensional parameters. Therefore, given the same value for the reaction order of secondary nucleation  $n_2$ , all realizations of protein aggregation become identical upon an appropriate rescaling of the time variable, irrespective of the specific experimental conditions or the specific molecular details of the protein system under consideration. This arises since all such realizations are in the same basin of attraction of the fixed point indexed by this particular value of  $n_2$ , and thus fall into the same universality class, regardless of their individual values for  $n_c$ ,  $k_2$ , etc. In Fig. 3(a) we have tested this prediction under different conditions on a range of unrelated protein aggregation systems, ranging from infectious prion systems to amyloidogenic proteins: bovine insulin [17], the WW domain [21], the yeast prion Ure2p [22], and IAPP [27]. We find that, upon appropriate rescaling, experimental aggregation data recorded at different initial monomer concentrations for WW, Ure2p, and insulin collapse onto a single universal curve described by Eq. (17) with  $n_2 = 0$  and corresponding to systems dominated by filament fragmentation. The aggregation of IAPP is driven

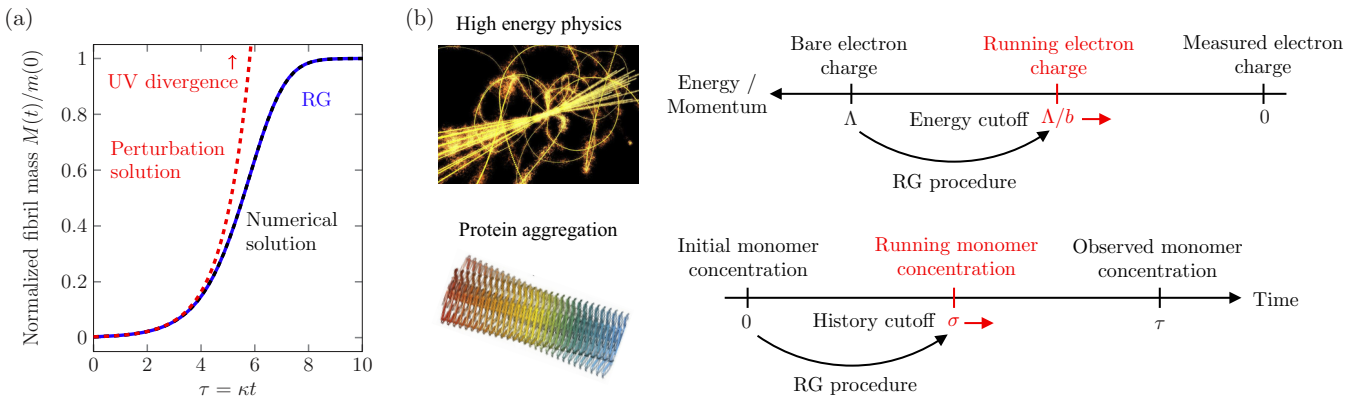


FIG. 2. (a) Comparison of the perturbative solution (6) (dashed red curve) and the RG solution (15) (solid blue curve) with the numerical solution to Eqs. (1) (dashed black curve) for  $n_2 = 0$  and  $\varepsilon = \lambda^2/2\kappa^2 = 3 \times 10^{-3}$ . (b) Schematics of the RG approach to filamentous aggregation kinetics and analogy to the RG procedure for high-energy physics. A running time variable  $\sigma$  connects the observed monomer concentration at time  $\tau = e^{\kappa t}$  with the known initial monomer concentration. The RG procedure (12) describes how the initial monomer concentration  $\rho$  is renormalized as the renormalization scale  $\sigma$  runs between the initial time and  $\tau$ . The amyloid structure figure is extracted from [41].



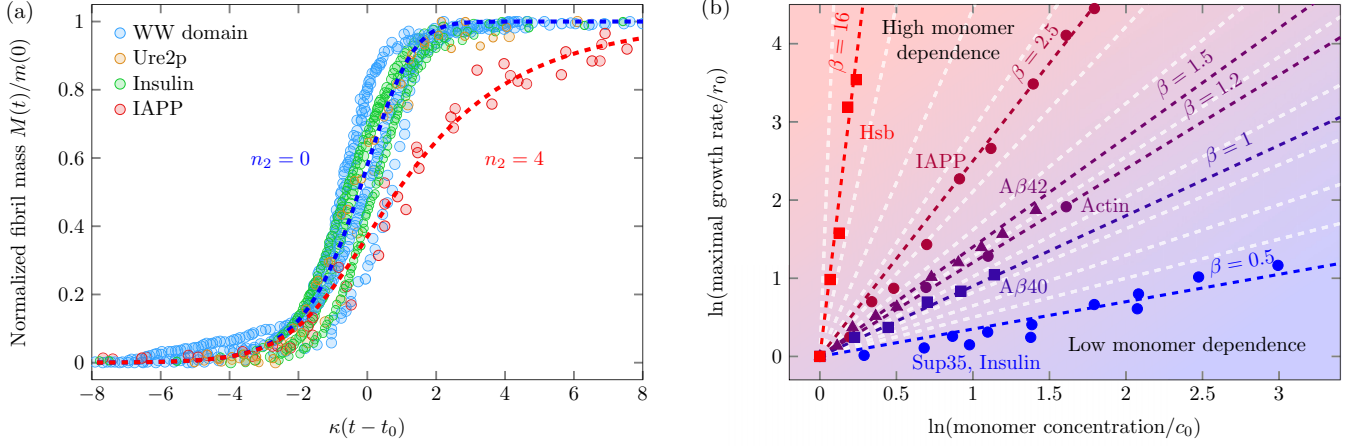


FIG. 3. Universality of protein filament formation kinetics. (a) Experimental data of filament mass formation at different initial monomer concentrations for insulin (2, 4, and 6  $\mu\text{M}$ ) [17], the WW domain (50, 100, 200, and 500  $\mu\text{M}$ ) [21], and the yeast prion Ure2p (20, 25, and 38  $\mu\text{M}$ ) [22] collapse onto a single universal curve upon appropriate rescaling of the time coordinate, where  $t_0 = -1/\kappa \ln(\varepsilon)$ . The universal curve (blue dashed line) is Eq. (17) with  $n_2 = 0$ . Upon rescaling, the aggregation kinetics data for IAPP [27] recorded at three different initial monomer concentrations (700, 800, and 1000  $\mu\text{M}$ ) collapse onto a different universal curve, Eq. (17) with  $n_2 = 4$  (red dashed line). (b) Classification of different filamentous protein systems into universality classes on the basis of the scaling exponent  $\beta$  of maximal growth rate. In this plot, universality classes corresponding to a different reaction order for secondary nucleation  $n_2$  (monomer dependence) are described by straight lines with slope  $\beta = (n_2 + 1)/2$  (shaded dashed lines). Data are shown for Sup35 and insulin ( $n_2 = 0$ ) [17], A $\beta$ 40 ( $n_2 = 1$ ) [31], actin in the presence of the Arp2/3 complex, activated by the WASP protein ( $n_2 = 1.1$ ) [25], A $\beta$ 42 ( $n_2 = 2$ ), IAPP ( $n_2 = 4$ ) [27], and sickle-cell hemoglobin ( $n_2 = 31$ ) [26]. In (b), each data set has been rescaled by the data point with the lowest monomer concentration;  $c_0$  thus indicates the lowest value of monomer concentration in each data set and  $r_0$  indicates the respective value for the maximal growth rate.

by surface catalyzed secondary nucleation with  $n_2 = 4$  and thus, as expected, experimental data of IAPP fibril formation at varying initial monomer concentrations collapse but onto a different universal curve.

### B. Scaling behavior

The RG also provides insights into the scaling behavior of protein aggregation [17,27,40]. Indeed, a direct implication of the RG is not only that filamentous assembly displays scaling behavior, but most importantly that the characteristic scaling exponents must depend solely on the universality class. This prediction is verified directly for the scaling of the reaction half-time  $t_{1/2}$ , which, from (17), is found to scale as  $t_{1/2} \propto m_{\text{tot}}^\gamma$ , where the exponent is

$$\gamma = -\frac{n_2 + 1}{2}. \quad (19)$$

Another important example is the scaling of the maximal growth rate  $r_{\text{max}} = \max[dM(t)/dt]$  with initial monomer concentration. In this case,  $r_{\text{max}} \propto m_{\text{tot}}^\beta$ , where the scaling exponent

$$\beta = \frac{n_2 + 1}{2} \quad (20)$$

is once again solely dependent on  $n_2$ .

### C. Classification of protein aggregation systems into universality classes

Since elements of the same universality class share the same scaling behavior, scaling exponents are the most direct way to classify different protein aggregation systems

into universality classes using experimental data, as shown in Fig. 3(b). Notably, the different universality classes that we find for protein aggregation can be related to simple autocatalytic processes. Indeed, the universal curves (17) are solutions to the RG equation, which is found to be equivalent to the generalized logistic equation

$$\frac{dM(t)}{dt} = \theta \kappa \{1 - [M(t)/K]^{1/\theta}\} M(t), \quad (21)$$

with  $K = m_{\text{tot}}$ . Generalized logistic equations represent the simplest description of autocatalytic growth with finite resources and thus emerge most commonly in population dynamics, where  $K$  is known as the carrying capacity, i.e., the maximum population size that the environment can sustain indefinitely, given the available energy resources, e.g., in the form of nutrients. For protein aggregation, we find that  $K$  is directly related to the chemical potential of monomers  $k_B T \ln(m_{\text{tot}})$ , i.e., the thermodynamic driving force for self-assembly. Thus, we may interpret monomer concentration as the available energy resources and we can interpret the reaction order  $n_2$ , which labels our different universality classes, as describing the dependence of fibril proliferation on these resources. Protein systems with  $n_2 < 1$ , such as insulin or the prion protein Sup35, have small  $1/\theta$  and are characterized by a weak dependence of proliferation on the available resources such that their growth curves respond only gently to changes in the monomer chemical potential. In the limiting case of fragmentation ( $n_2 = 0$ ), there is no resource dependence for the secondary process responsible for autocatalysis, and resource dependence is almost solely due to fibril elongation. By contrast, the assembly of protein systems with  $n_2 > 1$  and large  $1/\theta$ , such as A $\beta$ 42 or sickle-cell hemoglobin, is strongly

dependent on the available resources such that the approach of  $M(t)$  to the carrying capacity is much more gradual as the remaining resources dwindle. The growth of systems with  $n_2 = 1$ , such as A $\beta$ 40 in the concentration range 5–10  $\mu$ M, is equally as limited by resource depletion as it is accelerated by the accumulation of substrate the resources are converted into and is described by a logistic universal function.

Systems where the production of new aggregates is dominated by primary nucleation can be equally classified in universality classes in terms of the reaction order for primary nucleation  $n_c$  [1].

#### IV. SUMMARY AND OUTLOOK

By applying perturbative RG approaches to protein aggregation, we have found that the kinetics of these phenomena display universal behavior and that the resulting universality classes correspond to simple autocatalytic processes. The dependence of the rate of fibril self-replication on the available protein monomer resources provides a natural characterization of these universality classes. Looking forward, our work raises the natural question of whether, using

perturbative RG, the kinetics of general self-assembly systems could also be classified into universality classes similar to those found for filamentous systems. In this sense, the approach presented in this work could represent an important starting point for constructing a unified dynamical theory of self-assembly.

#### ACKNOWLEDGMENTS

We acknowledge support from Peterhouse, Cambridge (T.C.T.M.), the Swiss National Science Foundation (T.C.T.M.), the Schiff Foundation (A.J.D.), the Wellcome Trust (T.P.J.K.), the Cambridge Centre for Misfolding Diseases (T.P.J.K.), the BBSRC (T.P.J.K.), and the Frances and Augustus Newman Foundation (T.P.J.K.). The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (Grant No. FP7/2007-2013) through the ERC grant PhysProt (Agreement No. 337969). We thank Robert L. Jack (Cambridge) and Peter B. Bolhuis (Amsterdam) for useful discussions.

T.C.T.M. and A.J.D. contributed equally to this work.

#### APPENDIX: DERIVATION OF RENORMALIZED EXPANSION

In this Appendix we provide the details pertaining to the derivation of the renormalized expansion (10). We start from the naive perturbation solution (7), which we renormalize by multiplying it by the renormalization constant  $\rho(\sigma)$  [Eq. (8)]. This yields

$$\begin{aligned} \mu'(\tau - \sigma, \sigma) = & \rho_0(\sigma) - \varepsilon[(\tau - \sigma)\rho_0(\sigma) + \sigma\rho_0(\sigma) - \delta\rho_1(\sigma)] + \varepsilon^2 \left[ \frac{(\tau - \sigma)^2}{2c}\rho_0(\sigma) + \frac{(\tau - \sigma)\sigma}{c}\rho_0(\sigma) + \frac{\sigma^2}{2c}\rho_0(\sigma) \right. \\ & \left. - \delta_1(\sigma)(\tau - \sigma) - \delta\rho_1(\sigma)\sigma + \delta\rho_2(\sigma) \right] + \mathcal{R}. \end{aligned} \quad (\text{A1})$$

We now choose the counterterm  $\delta\rho_1(\sigma)$  to absorb the UV divergent terms in  $\sigma$  at order  $\varepsilon$ :

$$\delta\rho_1(\sigma) = \sigma\rho_0(\sigma). \quad (\text{A2})$$

Inserting (A2) in the second-order term in Eq. (A1), we obtain

$$\begin{aligned} \mu'(\tau - \sigma, \sigma) = & \rho_0(\sigma) - \varepsilon(\tau - \sigma)\rho_0(\sigma) + \varepsilon^2 \left[ \frac{(\tau - \sigma)^2}{2c}\rho_0(\sigma) + \frac{(\tau - \sigma)\sigma}{c}\rho_0(\sigma) + \frac{\sigma^2}{2c}\rho_0(\sigma) - (\tau - \sigma)\sigma\rho_0(\sigma) - \sigma^2\rho_0(\sigma) \right. \\ & \left. + \delta\rho_2(\sigma) \right] + \mathcal{R}. \end{aligned} \quad (\text{A3})$$

Choosing  $\delta\rho_2(\sigma)$  to absorb the UV divergent terms in  $\sigma$  at order  $\varepsilon^2$  yields

$$\delta\rho_2(\sigma) = \sigma^2\rho_0(\sigma) - \frac{\sigma^2}{2c}\rho_0(\sigma). \quad (\text{A4})$$

We thus obtain the renormalized expansion

$$\begin{aligned} \mu'(\tau - \sigma, \sigma) = & \rho_0(\sigma) - \varepsilon(\tau - \sigma)\rho_0(\sigma) + \varepsilon^2 \left[ \frac{(\tau - \sigma)^2}{2c}\rho_0(\sigma) + \frac{(\tau - \sigma)\sigma}{c}\rho_0(\sigma) - (\tau - \sigma)\sigma\rho_0(\sigma) \right] + \mathcal{R} \\ = & \rho_0(\sigma) \left\{ 1 - \varepsilon(\tau - \sigma) + \varepsilon^2 \left[ \frac{\tau^2 - \sigma^2}{2c} - \sigma(\tau - \sigma) \right] + \mathcal{R} \right\}, \end{aligned} \quad (\text{A5})$$

which is Eq. (10).

- [1] F. Oosawa and S. Asakura, *Thermodynamics of the Polymerization of Protein* (Academic, New York, 1975).
- [2] Y. Mimori-Kiyosue, *Cytoskeleton* **68**, 603 (2011).
- [3] F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.* **86**, 27 (2017).
- [4] C. M. Dobson, *Cold Spring Harb. Perspect. Biol.* **9**, a023648 (2017).
- [5] T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, *Nat. Rev. Mol. Cell Biol.* **15**, 384 (2014).
- [6] E. Gazit, *Chem. Soc. Rev.* **36**, 1263 (2007).
- [7] T. P. J. Knowles and R. Mezzenga, *Adv. Mater.* **28**, 6546 (2016).
- [8] G. Wei, Z. Su, N. P. Reynolds, P. Arosio, I. W. Hamley, E. Gazit, and R. Mezzenga, *Chem. Soc. Rev.* **46**, 4661 (2017).
- [9] C. M. Dobson, *Trends Biochem. Sci.* **24**, 329 (1999).
- [10] A. J. Baldwin *et al.*, *J. Am. Chem. Soc.* **133**, 14160 (2011).
- [11] K. G. Wilson, *Rev. Mod. Phys.* **47**, 773 (1975).
- [12] G. I. Barenblatt, *Scaling* (Cambridge University Press, Cambridge, 2003).
- [13] T. Kunihiro, *Prog. Theor. Phys.* **94**, 503 (1995).
- [14] N. D. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Addison-Wesley, Reading, 1992).
- [15] L.-Y. Chen, N. Goldenfeld, and Y. Oono, *Phys. Rev. Lett.* **73**, 1311 (1994).
- [16] L.-Y. Chen, N. Goldenfeld, and Y. Oono, *Phys. Rev. E* **54**, 376 (1996).
- [17] T. P. J. Knowles, C. A. Waudby, G. L. Devlin, S. I. A. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Wellend, and C. M. Dobson, *Science* **326**, 1533 (2009).
- [18] A. Wegner and P. Savko, *Biochemistry* **21**, 1909 (1982).
- [19] S. R. Collins, A. Douglass, R. D. Vale, and J. S. Weissman, *PLoS Biol.* **2**, e321 (2004).
- [20] W.-F. Xue, S. W. Homans, and S. E. Radford, *Proc. Natl. Acad. Sci. USA* **105**, 8926 (2008).
- [21] N. Ferguson, J. Berriman, M. Petrovich, T. D. Sharpe, J. T. Finch, and A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **100**, 9814 (2003).
- [22] L. Zhu, X.-J. Zhang, L.-Y. Wang, J.-M. Zhou, and S. Perrett, *J. Mol. Biol.* **328**, 235 (2003).
- [23] C. B. Andersen, H. Yagi, M. Manno, V. Martorana, T. Ban, G. Christiansen, D. E. Otzen, Y. Goto, and C. Rischel, *Biophys. J.* **96**, 1529 (2009).
- [24] K. J. Amann and T. D. Pollard, *Nat. Cell Biol.* **3**, 306 (2001).
- [25] H. N. Higgs, L. Blanchoin, and T. D. Pollard, *Biochemistry* **38**, 15212 (1999).
- [26] F. A. Ferrone, J. Hofrichter, and W. A. Eaton, *J. Mol. Biol.* **183**, 611 (1985).
- [27] A. M. Ruschak and A. D. Miranker, *Proc. Natl. Acad. Sci. USA* **104**, 12341 (2007).
- [28] S. I. A. Cohen, S. Linse, L. M. Luheshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, *Proc. Natl. Acad. Sci. USA* **110**, 9758 (2013).
- [29] A. Cacciuto, S. Auer, and D. Frenkel, *Nature (London)* **428**, 404 (2004).
- [30] A. Saric, A. K. Buell, G. Meisl, T. C. T. Michaels, C. M. Dobson, S. Linse, T. P. J. Knowles, and D. Frenkel, *Nat. Phys.* **12**, 874 (2016).
- [31] G. Meisl, X. Yang, E. Hellstrand, B. Frohm, J. B. Kirkegaard, S. I. A. Cohen, C. M. Dobson, S. Linse, and T. P. J. Knowles, *Proc. Natl. Acad. Sci. USA* **111**, 9384 (2014).
- [32] T. C. T. Michaels and T. P. J. Knowles, *Am. J. Phys.* **82**, 476 (2014).
- [33] T. C. T. Michaels, A. Saric, J. Habchi, S. Chia, G. Meisl, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, *Annu. Rev. Phys. Chem.* **69**, 273 (2018).
- [34] A. Lomakin, D. B. Teplow, D. A. Kirschner, and G. B. Benedek, *Proc. Natl. Acad. Sci. USA* **94**, 7942 (1997).
- [35] A. Saric, Y. C. Chebaro, T. P. J. Knowles, and D. Frenkel, *Proc. Natl. Acad. Sci. USA* **111**, 17869 (2014).
- [36] A. J. Dear, A. Saric, T. C. T. Michaels, C. M. Dobson, and T. P. J. Knowles, *J. Phys. Chem. B* **122**, 11721 (2018).
- [37] A. Saric, T. C. T. Michaels, A. Zacccone, T. P. J. Knowles, and D. Frenkel, *J. Chem. Phys.* **145**, 211926 (2016).
- [38] C. T. Lee and E. M. Terentjev, *J. Chem. Phys.* **147**, 105103 (2017).
- [39] T. C. T. Michaels, S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, *Phys. Rev. Lett.* **116**, 038101 (2016).
- [40] S. I. A. Cohen, M. Vendruscolo, M. E. Welland, C. M. Dobson, E. M. Terentjev, and T. P. J. Knowles, *J. Chem. Phys.* **135**, 065105 (2011).
- [41] L. Gremer *et al.*, *Science* **358**, 116 (2017).